

1 **Combining fine-scale social contact data with epidemic modelling**
2 **reveals interactions between contact tracing, quarantine, testing and**
3 **physical distancing for controlling COVID-19**

4
5 Josh A. Firth^{1,2}, Joel Hellewell³, Petra Klepac^{3,4}, Stephen Kisslers⁵, CMMID COVID-19 working
6 group, Adam J. Kucharski³, Lewis G. Spurgin^{6*}

7
8 1. Department of Zoology, University of Oxford, Oxford, UK

9 2. Merton College, University of Oxford, Oxford, UK

10 3. Centre for the Mathematical Modelling of Infectious Diseases, Department of Infectious Disease
11 Epidemiology, London School of Hygiene & Tropical Medicine, London, UK

12 4. Department for Applied Mathematics and Theoretical Physics, University of Cambridge

13 5. Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public
14 Health, Boston MA

15 6. School of Biological Sciences, University of East Anglia, Norwich, UK

16
17 ***Correspondence:** l.spurgin@uea.ac.uk

18
19 *CMMID COVID-19 working group members (order selected at random): Mark Jit, Katherine E.*

20 *Atkins, Samuel Clifford, C Julian Villabona-Arenas, Sophie R Meakin, Charlie Diamond, Nikos I*

21 *Bosse, James D Munday, Kiesha Prem, Anna M Foss, Emily S Nightingale, Kevin van Zandvoort,*

22 *Nicholas G. Davies, Hamish P Gibbs, Graham Medley, Amy Gimma, Stefan Flasche, David*

23 *Simons, Megan Auzenberg, Timothy W Russell, Billy J Quilty, Eleanor M Rees, Quentin J*

24 *Leclerc, W John Edmunds, Sebastian Funk, Rein M G J Houben, Gwenan M Knight, Sam Abbott,*

25 *Fiona Yueqian Sun, Rachel Lowe, Damien C Tully, Simon R Procter, Christopher I Jarvis, Akira*

26 *Endo, Kathleen O'Reilly, Jon C Emery, Thibaut Jombart, Alicia Rosello, Arminder K Deol, Matthew*

27 *Quaire, Stephanie Hue, Yang Liu, Rosalind M Eggo, Carl A B Pearson*

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

28 **Abstract**

29 Case isolation and contact tracing can contribute to the control of COVID-19 outbreaks^{1,2}.
30 However, it remains unclear how real-world networks could influence the effectiveness and
31 efficiency of such approaches. To address this issue, we simulated control strategies for SARS-
32 CoV-2 in a real-world social network generated from high resolution GPS data^{3,4}. We found that
33 tracing contacts-of-contacts reduced the size of simulated outbreaks more than tracing of only
34 contacts, but resulted in almost half of the local population being quarantined at a single point in
35 time. Testing and releasing non-infectious individuals led to increases in outbreak size, suggesting
36 that contact tracing and quarantine may be most effective when it acts as a 'local lockdown' when
37 contact rates are high. Finally, we estimated that combining physical distancing with contact
38 tracing could enable epidemic control while reducing the number of quarantined individuals. Our
39 approach highlights the importance of network structure and social dynamics in evaluating the
40 potential impact of SARS-CoV-2 control.

41 **Main**

42 Non-pharmaceutical interventions (NPIs) are central to reducing SARS-CoV-2 transmission^{5–8}.
43 Such responses generally include: case isolation, tracing and quarantining of contacts, use of PPE
44 and hygiene measures, and policies designed to encourage physical distancing (including closures
45 of schools and workplaces, banning of large public events and restrictions on travel). Due to the
46 varying economic and social costs of these interventions, there is a clear need for sustainable
47 strategies that limit SARS-CoV-2 transmission while reducing disruption as far as possible.

48
49 Isolation of symptomatic cases, and quarantine of their contacts (e.g. household members), is a
50 common public health strategy for reducing disease spread^{1,2,8}. This approach has been used as
51 part of SARS-CoV-2 control strategies⁹. However, the relatively high reproduction number of the
52 SARS-CoV2 virus in early outbreak stages^{10,11}, alongside likely high contribution to transmission
53 from presymptomatic and asymptomatic individuals¹², means that manual tracing of contacts alone
54 may not be a sufficient containment strategy under a range of outbreak scenarios¹³. As countries
55 relax lockdowns and other more stringent physical distancing measures, combining the isolation of
56 symptomatic individuals and quarantine of contacts identified through fine-scale tracing is likely to
57 play a major role in many national strategies for targeted SARS-CoV-2 control¹⁴.

58
59 It is possible to assess the potential effectiveness of contact tracing by simultaneously modelling
60 disease spread and contact tracing strategies through social systems of individuals¹⁵. These
61 systems are usually simulated through parameterisation with simple social behaviours (e.g. the
62 distribution of the number of physical contacts per individual). Further still, social systems may be
63 simulated as networks that can be parameterised according to assumptions regarding different
64 contexts (for example, with different simulated networks for households, schools and workplaces),
65 or using estimated contact rates of different age groups¹⁶. However, much less is known about
66 how different types of real-world social behaviour and the hidden structure found in real-life
67 networks may affect both patterns of disease transmission and efficacy of contact tracing under
68 different scenarios^{17,18}. Examining contagion dynamics and control strategies using a ‘real-world’

69 network allows for a more realistic simulation of SARS-CoV-2 outbreak and contact tracing
70 dynamics.

71

72 Here we develop an epidemic model which simulates COVID-19 outbreaks across a real-world
73 network, and we assess the impact of a range of testing and contact tracing strategies for
74 controlling these outbreaks (Table 1). We then simulate physical distancing strategies and quantify
75 how the interaction between physical distancing, contact tracing and testing affects outbreak
76 dynamics.

77

78 We used a publicly available dataset on human social interactions collected specifically for
79 modelling infectious disease dynamics as part of the British Broadcasting Corporation (BBC)
80 documentary “Contagion! The BBC Four Pandemic”. The high-resolution data collection focused
81 on residents of the town of Haslemere, where the first evidence of UK-acquired infection with
82 SARS-CoV-2 would later be reported in late February 2020¹⁹. Previous analyses of this dataset
83 have shown that it is structurally relevant to modelling disease spread, and hence holds substantial
84 potential for understanding and controlling real-world diseases⁴. Here, we defined dyadic contacts
85 on a day-by-day basis as at least one daily 5 min period with a distance of 4 m (see Methods),
86 which gave 1616 daily contact events and 1257 unique social links between 468 individuals. The
87 social network defined in this way was strongly correlated ($r > 0.85$ in all cases) with social
88 networks based on other contact distances (from 1-7 m contact ranges; Extended Data Fig. 1).
89 Similarly, social networks created using different time-periods for weighting the dyadic contacts
90 (Extended Data Fig. 2) were also strongly related to the weighting used here (i.e. number of days
91 seen together). As such, this social network quantification gives a representative indication of daily
92 contact propensities within the relevant transmission range between individuals (see Methods) and
93 also captures various aspects of the patterns and structure presented by different quantifications of
94 this social system.

95

96 Example outbreaks across the Haslemere social network under different control scenarios are
97 displayed in Fig. 1, with a full animated visualisation in Supplementary Video 1, and a Shiny app is
98 available to run individual outbreak simulations (see data sharing). Across all simulations, our
99 epidemic model showed that uncontrolled outbreaks in the Haslemere network stemming from a
100 single infected individual resulted in a median of 75% (IQR = 74%-76%) of the population infected
101 after 70 days (Fig. 2). Isolation when symptomatic resulted in 66% (65%-67%) of the population
102 infected, while primary contact tracing resulted in 48% (46%-50%) infected. Secondary contact
103 tracing resulted in the smallest percentage (16%, 14%-19%) of the population infected after 70
104 days. The number of quarantined individuals was very high under both primary and secondary
105 contact tracing, with a median of 43% (IQR = 32%-52%) of the population quarantined during the
106 outbreak peak with the latter (Fig. 2). Examining temporal dynamics showed that outbreak peaks
107 typically occurred within the first 1-3 weeks, and that interventions reduced the overall size of the
108 outbreaks as well as their growth rate (Fig. 2). The number of people required to isolate or
109 quarantine followed a similar trajectory to the number of cases, although under secondary contact
110 tracing, substantial proportions of the population (27%, 18%-35%) were quarantined even at the
111 end of the simulations (Fig. 2). This is in line with a large-scale recent simulation model of app-
112 based contact tracing in the UK₂₀, which suggested that contact tracing could be highly effective,
113 but also that it required large numbers of people to be quarantined. Further, in our (optimistic)
114 default parameter settings we assumed that 10% of contact tracing attempts were missed. This,
115 combined with the very large number of quarantined cases under secondary contact tracing (Fig.
116 2), suggests that a majority of the population could receive a notification that they should
117 quarantine within the first 2-3 weeks of an outbreak.

118

119 Sensitivity analysis of the efficacy of contact tracing under the epidemic model is presented in
120 Extended Data Figs 3-6. As expected, outbreak size decreased with the percentage of contacts
121 traced in all scenarios, and increased with the reproduction number, the proportion of
122 asymptomatic cases, the proportion of pre-onset transmission, the delay between onset/tracing
123 and isolation/quarantine, and the number of initial cases (Extended Data Figs 3-6). Outbreak

124 dynamics were strongly affected by outside infection rate across all intervention scenarios, as were
125 the number of isolated and quarantined cases (Extended Data Fig. 6). Our model therefore
126 corroborates with models using simulated social systems and showing that, for a disease such as
127 COVID-19 with high levels of transmission from asymptomatic and presymptomatic individuals,
128 contact tracing is likely to be most effective when the proportion of traced contacts is high, when
129 the delay from notification to quarantine is short¹³, and, most importantly, when the number of
130 starting cases and rate of movement into the network are low. Importantly, however, the tradeoff
131 between the number of cases and the number of quarantined cases was found across the entirety
132 of the parameter space (Extended Data Figs 3-6). Further, increasing the network density through
133 increasing the distance threshold for defining contacts led to broadly similar results across
134 intervention scenarios, albeit with larger numbers of quarantined cases required for outbreak
135 control via contact tracing (Extended Data Fig. 7). Therefore, while more real-world networks are
136 needed to demonstrate how well these results apply to other locations and settings, our results are
137 robust to a range of epidemiological and network parameters.

138

139 The number of quarantined cases can be reduced through mass testing and release of individuals
140 who return a negative result. Conversely, if contact rates in the population are high, large-scale
141 test and release strategies could provide greater opportunity for transmission and decrease the
142 effectiveness of contact tracing. We therefore assessed how the testing and releasing of isolated
143 and quarantined subjects might affect the numbers of cases and time spent in isolation and
144 quarantine, using false positive and false negative rates estimated from empirical data^{21,22}
145 (Supplementary Table 1). We estimated that increasing the testing capacity (and therefore testing
146 and releasing more quarantined cases) led to substantial increases in outbreak size, especially
147 under secondary contact tracing (median = 50%, IQR = 48%-52%; Fig. 3A). This result occurred
148 despite an optimistic false negative rate of 10%, suggesting that the increase in outbreak size with
149 high testing rates is a result of increased transmission within the network, rather than through
150 releasing infected cases *per se*. Therefore, secondary tracing may effectively act as a 'local
151 lockdown' rather than a targeted intervention strategy. High levels of testing did not lead to large

152 reductions in the number of quarantined cases under secondary contact tracing scenarios, and the
153 number of tests required to reduce the numbers of quarantined cases were large, with 68% (63%-
154 71%) of the population requiring tests in a single week during outbreak peaks (Fig. 3A). We cannot
155 be sure to what extent our results will represent larger populations, but the tripartite relationship
156 between the number of cases, the number of quarantined contacts and the number of tests
157 required will apply in the majority of scenarios in which rates of social interaction are high.

158

159 A very high notification and quarantine rate for any contact tracing system may have
160 consequences for adherence. Our model is optimistic in its assumption that individuals isolate
161 independently of previous notifications or isolations, and highly optimistic in its assumption of
162 100% adherence to quarantine among traced contacts. In reality a high notification and quarantine
163 rate may result in individuals being less likely to undertake quarantine in the future, which in turn
164 will impact outbreak dynamics. There is a need for more evidence and models to better understand
165 these behavioural dynamics, in order to develop sustainable intervention strategies²³. One
166 suggested solution to reduced adherence to quarantine is through (digital) targeted quarantine
167 requests to the individuals at highest risk of infection, or to those most likely to spread to others²⁴.
168 To what extent these interventions will be needed, and how well they will work, is not yet clear;
169 however, our study provides a methodological template for network-based research into SARS-
170 CoV2 and its potential control strategies.

171

172 Combining contact tracing with other physical distancing measures may allow for outbreak control
173 while reducing the number of people in quarantine, and the number of tests required. We
174 simulated physical distancing by reducing the number of weak links in the Haslemere network. We
175 aimed to consider low to moderate levels of physical distancing, so used a model whereby the only
176 interactions with 'rare' contacts are removed. We found that, across all scenarios, physical
177 distancing led to reductions in the number of overall cases (Fig. 3B). Importantly, increasing
178 physical distancing was associated with lower numbers of quarantined cases, which was reduced
179 to as little as 6% (3%-9%) during outbreak peaks under secondary contact tracing (Fig. 3B).

180 Simulating physical distancing using an alternative approach whereby removed ‘rare’ contacts
181 were reassigned to existing contacts (see methods) yielded similar results to our simpler model,
182 although using this approach, physical distancing led to smaller decreases in outbreak size
183 (Extended Data Fig. 8). We do not have information on household structure within the Haslemere
184 dataset, but our physical distancing scenario is analogous to decreasing the level of non-
185 household contacts. Therefore it may be that combining measures that reduce non-household
186 contact rates with highly effective contact tracing may be a useful tool for control of SARS-CoV-2
187 spread. However, further work is required to determine exactly what kinds of physical distancing
188 measures would enable effective outbreak control alongside contact tracing. Furthermore, future
189 investigations could also examine how the spread of the disease itself may shape behavioural
190 change interventions (e.g. where large outbreaks spark more severe physical distancing
191 measures), and how this feedback may shape the contagion dynamics and predicted effectiveness
192 of interventions.

193

194 Network structure can have substantial effects on epidemic model predictions^{25,26}. To investigate
195 this, we used null network models based on the Haslemere data, which maintained the same
196 number of individuals, connections and weights of connections, but shuffled network architecture in
197 different ways (see Methods). The number of cases estimated using the null networks was broadly
198 similar to the real-world network, although this was substantially underestimated in a ‘lattice’ like
199 network (Fig. 4). Importantly, the rate of quarantine varied substantially among the null networks,
200 especially under secondary contact tracing (Fig. 4). These results demonstrate that the use of
201 network-based simulations of SARS-CoV-2 dynamics requires caution, as even if such models had
202 precise information on the number of individuals and amount of social interactions occurring within
203 a system, the assumed architecture of the social network structure alone can shape predictions for
204 both the extent of spread and the usefulness of control strategies. Furthermore, through providing
205 insight into how changes to network structure influences contagion dynamics, the null network
206 simulation approach gives some indication of how this contagion and associated control strategies
207 may operate in different social environments. For instance, different social structures may arise

208 when considering particular social settings (e.g. workplaces, commuting), some of which may be
209 closer to the null networks generated here. Considering this structure will lead to improved
210 predictions of outbreak dynamics.

211

212 There are a number of important limitations to our study and the current availability of empirical
213 data in general. Most importantly, this social network is taken from a single, small town and over a
214 short period of time, and we do not know to what extent the social dynamics will be applicable to
215 larger cities and other contexts and over long periods. Therefore, future large-scale efforts in
216 gathering data on dynamic fine-scale social behaviour over long periods of time (ideally over the
217 entire contagion period) in major cities would be of great benefit for assessing the relative uses of
218 SARS-CoV-2 control strategies, and for understanding how and why interventions implemented in
219 some cities have been relatively more successful than others²⁷. Indeed, the epidemic network-
220 based model provided here can be applied generally to more extensive social networks if such
221 data becomes available in the future. Further, the Haslemere data, while rich, does not sample the
222 entire population of Haslemere, and children under the age of 13 were not included in the
223 experiment, which could potentially have an impact on outbreak and social tracking dynamics.
224 Again, such issues are also likely to be prevalent across real-world contact-tracing attempts, as the
225 ability to track children will be limited, particularly with app-based approaches that require a
226 smartphone. It is encouraging that our results broadly align with other, larger-scale simulations of
227 contact tracing which explicitly model these limitations, but lack the fine-scale social tracking
228 data²⁰. Therefore, by supplying a general framework for simulating the spread of COVID-19 on
229 real-world networks, we hope to promote integration of multiple real-world social tracking datasets
230 with epidemic modelling, which may provide a promising way forward for optimising contact tracing
231 strategies and other non-pharmaceutical interventions.

232

233 **References**

- 234 1. Fraser, C., Riley, S., Anderson, R. M. & Ferguson, N. M. Factors that make an infectious
235 disease outbreak controllable. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 6146–6151 (2004).

- 236 2. Peak, C. M., Childs, L. M., Grad, Y. H. & Buckee, C. O. Comparing nonpharmaceutical
237 interventions for containing emerging epidemics. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 4023–
238 4028 (2017).
- 239 3. Kissler, S. M., Klepac, P., Tang, M., Conlan, A. J. K. & Gog, J. R. Sparking ‘The BBC Four
240 Pandemic’: Leveraging citizen science and mobile phones to model the spread of disease.
241 *bioRxiv* 479154 (2018) doi:10.1101/479154.
- 242 4. Klepac, P., Kissler, S. & Gog, J. Contagion! The BBC Four Pandemic--The model behind the
243 documentary. *Epidemics* **24**, 49–59 (2018).
- 244 5. Ferguson, N. *et al.* Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce
245 COVID19 mortality and healthcare demand. (2020).
- 246 6. Chinazzi, M. *et al.* The effect of travel restrictions on the spread of the 2019 novel coronavirus
247 (COVID-19) outbreak. *Science* **368**, 395–400 (2020).
- 248 7. Tian, H. *et al.* An investigation of transmission control measures during the first 50 days of the
249 COVID-19 epidemic in China. *Science* (2020) doi:10.1126/science.abb6105.
- 250 8. Aleta, A. *et al.* Modeling the impact of social distancing, testing, contact tracing and household
251 quarantine on second-wave scenarios of the COVID-19 epidemic. *medRxiv* (2020)
252 doi:10.1101/2020.05.06.20092841.
- 253 9. Chen, S. What’s behind Vietnam’s coronavirus containment success? *South China Morning*
254 *Post* [https://www.scmp.com/news/asia/southeast-asia/article/3079598/coronavirus-whats-](https://www.scmp.com/news/asia/southeast-asia/article/3079598/coronavirus-whats-behind-vietnams-containment-success)
255 [behind-vietnams-containment-success](https://www.scmp.com/news/asia/southeast-asia/article/3079598/coronavirus-whats-behind-vietnams-containment-success) (2020).
- 256 10. Kucharski, A. J. *et al.* Early dynamics of transmission and control of COVID-19: a
257 mathematical modelling study. *Lancet Infect. Dis.* (2020) doi:10.1016/S1473-3099(20)30144-
258 4.
- 259 11. Klinkenberg, D., Fraser, C. & Heesterbeek, H. The effectiveness of contact tracing in
260 emerging epidemics. *PLoS One* **1**, e12 (2006).
- 261 12. He, X. *et al.* Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat. Med.*
262 (2020) doi:10.1038/s41591-020-0869-5.
- 263 13. Hellewell, J. *et al.* Feasibility of controlling COVID-19 outbreaks by isolation of cases and

- 264 contacts. *Lancet Glob Health* **8**, e488–e496 (2020).
- 265 14. Ferretti, L. *et al.* Quantifying SARS-CoV-2 transmission suggests epidemic control with digital
266 contact tracing. *Science* (2020) doi:10.1126/science.abb6936.
- 267 15. Eames, K. T. D. & Keeling, M. J. Contact tracing and disease control. *Proc. Biol. Sci.* **270**,
268 2565–2571 (2003).
- 269 16. Del Valle, S. Y., Hyman, J. M., Hethcote, H. W. & Eubank, S. G. Mixing patterns between age
270 groups in social networks. *Soc. Networks* **29**, 539–554 (2007).
- 271 17. Kiss, I. Z., Green, D. M. & Kao, R. R. Disease contact tracing in random and clustered
272 networks. *Proc. Biol. Sci.* **272**, 1407–1414 (2005).
- 273 18. Read, J. M., Eames, K. T. D. & Edmunds, W. J. Dynamic social networks and the implications
274 for the spread of infectious disease. *J. R. Soc. Interface* **5**, 1001–1007 (2008).
- 275 19. BBC News. Coronavirus patient first to be infected in UK. *BBC* (2020).
- 276 20. Hinch, R. *et al.* *Effective Configurations of a Digital Contact Tracing App: A report to NHSX*.
277 https://github.com/BDI-pathogens/covid-19_instant_tracing (2020).
- 278 21. Chau, N. V. V. *et al.* The natural history and transmission potential of asymptomatic SARS-
279 CoV-2 infection. *Infectious Diseases (except HIV/AIDS)* (2020)
280 doi:10.1101/2020.04.27.20082347.
- 281 22. Cohen, A. N. & Kessel, B. False positives in reverse transcription PCR testing for SARS-CoV-
282 2. *Epidemiology* (2020) doi:10.1101/2020.04.26.20080911.
- 283 23. West, R., Michie, S., Rubin, G. J. & Amlôt, R. Applying principles of behaviour change to
284 reduce SARS-CoV-2 transmission. *Nat Hum Behav* **4**, 451–459 (2020).
- 285 24. McCall, B. Shut down and reboot-preparing to minimise infection in a post-COVID-19 era.
286 *Lancet Digit Health* (2020) doi:10.1016/S2589-7500(20)30103-5.
- 287 25. Keeling, M. J. & Eames, K. T. D. Networks and epidemic models. *J. R. Soc. Interface* **2**, 295–
288 307 (2005).
- 289 26. Xu, Z. & Sui, D. Z. Effect of Small-World Networks on Epidemic Propagation and Intervention.
290 *Geogr. Anal.* **41**, 263–282 (2009).
- 291 27. Cohen, J. & Kupferschmidt, K. Countries test tactics in ‘war’ against COVID-19. *Science* **367**,

- 292 1287–1288 (2020).
- 293 28. Cairns, S. J. & Schwager, S. J. A comparison of association indices. *Anim. Behav.* **35**, 1454–
294 1469 (1987).
- 295 29. Maslov, S. & Sneppen, K. Specificity and stability in topology of protein networks. *Science*
296 **296**, 910–913 (2002).
- 297 30. Davies, N. G. *et al.* The effect of non-pharmaceutical interventions on COVID-19 cases,
298 deaths and demand for hospital services in the UK: a modelling study. *Infectious Diseases*
299 *(except HIV/AIDS)* (2020) doi:10.1101/2020.04.01.20049908.
- 300

301 **Methods**

302

303 *Social tracking data*

304 The Haslemere dataset was generated and described as part of previous work, which gives
305 detailed description of the characteristics of this dataset and town^{3,4}. Briefly, the data were
306 collected during the 2017/18 *BBC Pandemic* project conducted in Haslemere, Surrey, UK. The
307 project involved a massive citizen-science experiment to collect social contact and movement data
308 using a custom-made phone app, and was designed to generate data relevant to understanding
309 directly transmitted infectious disease^{3,4}. Of the 1272 individuals within Haslemere that
310 downloaded the app, 468 individuals had sufficient data points at a resolution of 1m over three full
311 days within the focal area for further analysis³. All 468 focal individuals were known to have spent
312 >6hrs within 51.0132;-0.7731SW : 51.1195,-0.6432NE (within Postcode GU27), but the dataset
313 used here comprises of de-identified proximity data made available as pairwise distances (~1 m
314 resolution) at 5 min intervals (excluding 11pm-7am)³.

315

316 *Social network construction*

317 In our primary analysis, we defined social contacts as events when the average pairwise distances
318 between individuals within a 5 min time interval (calculated using the Haversine formula for great-
319 circle geographic distances³) are 4 m or less. By doing so, we aimed to capture the majority of
320 relevant face-to-face contacts (i.e. those that may result in transmission) over 5 min periods,
321 particularly given the 1 m potential error³ on the tracking measurement during these short time
322 intervals. Furthermore, this 4 m threshold is within typical mobile phone Bluetooth ranges for
323 relatively accurate and reliable detections. Therefore, this contact dataset will also be comparable
324 to proximity-based contacts identified through Bluetooth contact tracing apps, which may be
325 preferred to real-location tracking for privacy reasons. We considered the sensitivity of the network
326 to the contact definition by testing six further social networks from contacts defined using different
327 threshold distances spanning the conceivable potential transmission range within the 5 min
328 intervals (1 m to 7 m thresholds). We first measured the correlation of the network structure (i.e.

329 pairwise contacts) across the seven networks using Mantel tests. We also measured the
330 correlation of each individual's degree (number of contacts), clustering coefficient (number of
331 contacts also connected to one another), betweenness (number of shortest paths between nodes
332 that pass through an individual), and eigenvector centrality (a measure that accounts both for a
333 node's centrality and that of its neighbours) across the seven networks.

334

335 The Haslemere data is a temporal dataset spanning three full days. While the epidemic model we
336 use is dynamic (see below Methods), the contagion process of COVID-19 operates over a longer
337 time period than three days. To be able to meaningfully simulate longer-term outbreak dynamics,
338 we quantified the data as a static social network in which edges indicate the propensities for social
339 contact between nodes. Temporal information is incorporated by weighting the edges using the
340 temporal contact information, instead of using a dynamic network which would require contact data
341 over a much longer period. In the primary analysis, we weighted the edges as the number of
342 unique days a dyad was observed together (but see Supplementary Information for other temporal
343 definitions). Therefore, the weight score indicates the propensity for each dyad to engage in a
344 social contact event on any given day, whereby 0 = no contact, 1 = 'weak links' observed on the
345 minority of days (one third), 2 = 'moderate links' observed on the majority of days (two thirds), and
346 3 = 'strong links' observed on all days. In this way, the weights of this social network could be
347 included directly, and intuitively, into the dynamic epidemic model (see below). For sensitivity
348 analysis, we also created other weightings for this network, and examined the correlation in dyadic
349 social associations scores (using Mantel tests) with our primary weighting method (described
350 above). Specifically, for the sensitivity analysis, we used edges specified as i) a binary (i.e.
351 unweighted) network across all days, ii) a raw (and ranked) count of 5 min intervals in contact, iii) a
352 transformed weighted count (edge weight transformed as $1 - e^{-interval\ count}$, which approximates
353 a scenario where infection risk increases with contact time, but begins to level off after ~30mins of
354 contact between dyads) and iv) a 'simple ratio index' (SRI) weighting that corrects for observation
355 number as SRI score₂₈. The SRI score for any two individuals (i.e. A and B) is calculated as:

356

357 (1)
$$SRI_{A,B} = \frac{Obs_{A,B}}{Obs_A + Obs_B - Obs_{A,B}},$$

358

359 where *Obs* is the number of 5 min observation periods (the intervals since the start of the day)
360 within which an individual is recorded within 4 m of another individual.

361

362 *Null network simulation approach*

363 We used null networks²⁹ to understand the network properties that shape predictions of COVID-19
364 spread under different control scenarios. Null networks can also show how contagion may depend
365 on the arrangement of social ties, how it may operate in different social environments, and which
366 simulation approaches may be the most similar to real-world infection dynamics. We created four
367 null network scenarios (Extended Data Fig. 9) with 1000 networks generated under each of these.
368 All of the null network scenarios kept the same number of nodes, edges, and weights of these
369 edges, as the Haslemere network, but were generated under the following nulls: (1) 'edge null'
370 (Extended Data Fig. 9A) considered random social associates, allowing the edges of the network
371 to be randomly allocated between all nodes; (2) 'degree null' (Extended Data Fig. 9B) considered
372 individual differences in sociality but random social links between dyads, so randomly swapped the
373 edges between nodes but maintained the degree distribution of the real network (and was,
374 therefore, even more conservative than a power-law network simulation aiming to match real
375 differences in sociality); (3) 'lattice null' (Extended Data Fig. 9C) considered triadic and tight clique
376 associations, so created a ring-like lattice structure through assigning all edges into a ring-lattice
377 where individuals are connected to their direct neighbours, and their neighbours of the second and
378 third order (i.e. six links per individual), and then we randomly removed excess links (until the
379 observed number of edges was reached); (4) 'cluster null' (Extended Data Fig. 9D) considered the
380 observed level of clustering, so created a ring-lattice structure as described above but only
381 between individuals observed as connected (at least 1 social link) in the real network, added
382 remaining links (sampled from 4th order neighbours), and then rewired the edges until the real-
383 world global clustering was observed (~20% rewiring; Extended Data Fig. 9D). These conservative
384 (and informed) null models allowed connections to be arranged differently within the network but

385 maintained the exact same number of individuals, social connections and weights of these social
386 connections at each simulation.

387

388 *Epidemic model*

389 Building on the epidemiological structure of a previous branching-process model¹³, we developed a
390 full epidemic model to simulate COVID-19 dynamics across the Haslemere network. Full model
391 parameters are given in Supplementary Table 1. For a given network of individuals, an outbreak is
392 seeded by randomly infecting a given number of individuals (default = 1). The model then moves
393 through daily time steps, with opportunities for infection on each day. All newly infected individuals
394 are assigned an ‘onset time’ drawn from a Weibull distribution (mean = 5.8 days) that determines
395 the point of symptom onset (for symptomatic individuals), and the point at which infectiousness is
396 highest (for all individuals)¹². Each individual is then simultaneously assigned asymptomatic status
397 (whether they will develop symptoms at their onset time), as well as presymptomatic status
398 (whether or not they will infect before their assigned onset time), drawn from Bernoulli distributions
399 with defined probabilities (defaults = 0.4 and 0.2 respectively, see Supplementary Table 1). At the
400 start of each day, individuals are assigned a status of susceptible, infectious or recovered (which
401 would include deaths) based on their exposure time, onset time and recovery time (calculated as
402 onset time plus seven days), and are isolated or quarantined based on their isolation/quarantine
403 time (described below). The model then simulates infection dynamics over 70 days.

404

405 Possible infectors are all non-isolated and non-quarantined infectious individuals. Each day, all
406 susceptible, non-isolated, non-quarantined contacts of all infectors within the network are at risk of
407 being infected. The transmission rate for a given pair of contacts is modeled as:

408

$$409 \quad (2) \quad \lambda(t, s_i, p_i) = A_{s_i} I_{e_i} \int_{t-1}^t f(u; \mu_i, \alpha_{p_i}, \omega_{p_i}) du$$

410

411 where t is the number of days since the infector i was exposed, s_i and p_i are the infector’s
412 symptom status (asymptomatic yes/no, and presymptomatic yes/no, respectively). A_{s_i} is the scaling

413 factor for the infector's symptomatic status (Supplementary Table 1) and I_{ei} is the weighting of the
414 edge in the network (i.e. number of days observed together) between the infector and the
415 susceptible individual. The probability density function $f(u; \mu_i, \alpha_{p_i}, \omega_{p_i})$ corresponds to the
416 generation time, which is drawn from a skewed normal distribution (see 13 for details). Briefly, this
417 uses the infector's onset time as the location parameter μ_i , while the slant parameter α_{p_i} and the
418 scale parameter ω_{p_i} both vary according to the infector's presymptomatic transmission status
419 (Supplementary Table 1). This enabled us to simulate a predefined rate of presymptomatic
420 transmission, while retaining a correlation structure between onset time and infectiousness, and
421 avoiding a scenario whereby a large number of individuals were highly infectious on the first day of
422 exposure (see Supplementary Table 1 and data sharing for more details).

423
424 Using this transmission rate, the probability of infection between a susceptible-infected pair of
425 individuals t days after the infector's exposure time is then modeled as:

426
427 (3)
$$P(t, s_i, p_i) = 1 - e^{-\lambda(t, s_i, p_i)}$$

428
429 Note that the change in status from "infectious" to "recovered" at seven days after symptom onset
430 does not affect infection dynamics (as transmission rate ≈ 0 seven days after onset time in our
431 model), but is instead used for contact tracing purposes (see below). To test how the above rate of
432 infection related to the reproduction number R_0 and the observed generation times, we generated
433 empirical estimates of the number of secondary infections in the early outbreak stages of the
434 model. We ran 1000 trial simulations from a random single starting infector and quantified i) the
435 mean number of secondary infections from this case, and ii) the time at which each secondary
436 case was infected. We multiplied the rate of infection by a scaling parameter to get a baseline R_0 of
437 2.8, although we also performed sensitivity analysis (Supplementary Table 1). The mean
438 generation time using this method was 6.3 days (median = 6 days). These basic parameters
439 correspond closely to published estimates^{12,30}.

440

441 In addition to the infection rate from within the network, the infection rate from outside the network
442 is also simulated daily by randomly infecting susceptible individuals with a probability of 0.001
443 (although we also performed sensitivity analysis of this parameter).

444

445 We simulated different contact tracing scenarios using contact information from the network, with
446 the aim of evaluating both app-based and manual contact tracing strategies. Primary and
447 secondary contacts of individuals are identified from the network on the day of the infector's
448 symptom onset and, as such, contacts of asymptomatic infectors are not traced. Contacts who
449 have already recovered are excluded. Susceptible contacts are traced with a given probability (0.3-
450 0.9 tested - see Supplementary Table 1). We assume that this probability captures a wide range of
451 reasons why contacts might not be traced, and it thus acts as an intuitive simplification.

452

453 The isolation and/or quarantine time of each individual is determined based on their infection
454 status, their symptomatic status, whether they have been traced, and the control scenario. We
455 consider four control scenarios: i) no control, where no individuals are isolated or quarantined, ii)
456 case isolation, where individuals isolate upon symptom onset after a delay period, iii) primary
457 contact tracing with quarantine, where individuals isolate upon symptom onset (after a delay) and
458 traced contacts are quarantined upon their infector's symptom onset (also after a delay), and iv)
459 secondary contact tracing, as scenario iii) but including contacts of contacts. All isolated and
460 quarantined individuals are contained for 14 days.

461

462 Finally, we simulated a range of testing efforts for SARS-CoV-2. Each individual is assigned a
463 testing time on isolation or quarantine, with the delay between containment and testing sampled
464 from a Weibull distribution. A cap of the maximum number of daily tests is assigned, and each day
465 up to this number of individuals are randomly selected for testing. Test results are dependent on
466 infection and asymptomatic status, with a false negative rate (i.e. the probability that an infectious
467 case will test negative) 0.1²¹, and a false positive rate (i.e. the probability that susceptible case will

468 test positive) of 0.02₂₂. Cases who tested negative were immediately released from
469 isolation/quarantine.

470

471 A set of default parameters were chosen to represent a relatively optimistic model of contact
472 tracing, which included a short time delay between symptom onset/tracing and isolation/quarantine
473 (1-2 days), and a high proportion (90%) of contacts traced within this tracked population (default
474 parameters highlighted in bold in Supplementary Table 1). We assumed that the probability of
475 tracing was constant over time, and therefore independent of previous isolation/quarantine events,
476 and that all individuals remained in quarantine for the full 14 days, unless released via testing. We
477 performed sensitivity tests on all relevant parameters (Supplementary Table 1). To examine how
478 infection dynamics were affected by network structure, we ran epidemic simulations on each of the
479 null networks described above. We also ran simulations on networks generated using higher (7m
480 and 16m) distance thresholds for defining a contact. These networks were 20% and 100% more
481 dense, respectively, and therefore provide an estimate of the robustness of our simulations to
482 missing contacts.

483

484 We ran each simulation for 70 days, at which point the majority of new infections came from
485 outside the network (see results), with all scenarios replicated 1000 times. With the null networks
486 (above) and physical distancing simulations (below), we ran one replicate simulation on each of
487 1000 simulated networks. In no simulations were all individuals in the population infected under our
488 default settings. Therefore, for each simulation we report the number of cases per week, and
489 quantify the total number of cases after 70 days as a measure of outbreak severity. To present the
490 level of isolation and quarantine required under different scenarios, we calculate the number of
491 people contained on each day of the outbreak, and average this over weeks to get weekly changes
492 in the daily rates of isolations and quarantines.

493

494 *Physical distancing Simulations*

495 We simulated a population-level physical distancing effort, whereby a given proportion of the ‘weak
496 links’ are removed (edges only observed on a single day; Extended Data Fig. 10A-D). This is akin
497 to a simple situation whereby individuals reduce their non-regular contacts (e.g. to people outside
498 of their household or other frequently visited settings such as workplaces). As further
499 supplementary analysis, we also carried out a more complex physical distancing simulation,
500 whereby the weak links that were removed were randomly reassigned to existing contacts
501 (Extended Data Fig. 10E-G). This represents a scenario where individuals reduce their non-regular
502 contacts but spend more time with regular contacts.

503

504 The epidemic model code can be accessed at: <https://github.com/biouea/covidhm>

505

506 **Acknowledgements**

507 This work was instigated through the Royal Society’s Rapid Assistance in Modelling the Pandemic
508 (RAMP) scheme. We thank Michael Pointer for helpful discussions throughout, and Cock van
509 Oosterhout and Julia Gog for comments on the manuscript. We thank all those in Haslemere who
510 took part in the BBC Pandemic study. We thank Hannah Fry and 360 Production, especially
511 Danielle Peck and Cressida Kinnear, for making possible the collection of the dataset that
512 underlies this work, and Andrew Conlan, Maria Tang, and Julia Gog for their contribution to the
513 BBC study. JAF was supported by a research fellowship from Merton College and BBSRC
514 (BB/S009752/1) and acknowledges funding from NERC (NE/S010335/1). PK was in part funded by
515 the Royal Society under award RPEA180004 and European Commission: 101003688. AJK was
516 supported by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal
517 Society (grant Number 206250/Z/17/Z).

518

519 **Data availability**

520 This study used the raw data previously published in Kissler et al. 2018 (made available with full
521 description here: <https://www.biorxiv.org/content/10.1101/479154v1>). The data used here are
522 publicly available with the code

523

524 **Code availability**

525 The code and data used to produce the simulations is available as an R package at:

526 <https://github.com/biouea/covidhm>. A shiny app which runs individual outbreak simulations is

527 available at: https://biouea.shinyapps.io/covidhm_shiny/

528

529 **References**

530

531 **Author contributions**

532 J.A.F. A.J.K. and L.G.S. conceived the study; J.A.F. carried out the social network analysis, with

533 input from P.K., S.K., A.J.K and L.G.S; L.G.S. built the epidemic network model with input from

534 J.A.F., J.H., S.K., P.K. and A.J.K; J.A.F. and L.G.S. wrote the first draft of the manuscript; All

535 authors interpreted the results, contributed to writing and approved the final version for submission.

536

537 **Competing interests**

538 The authors declare no competing interests

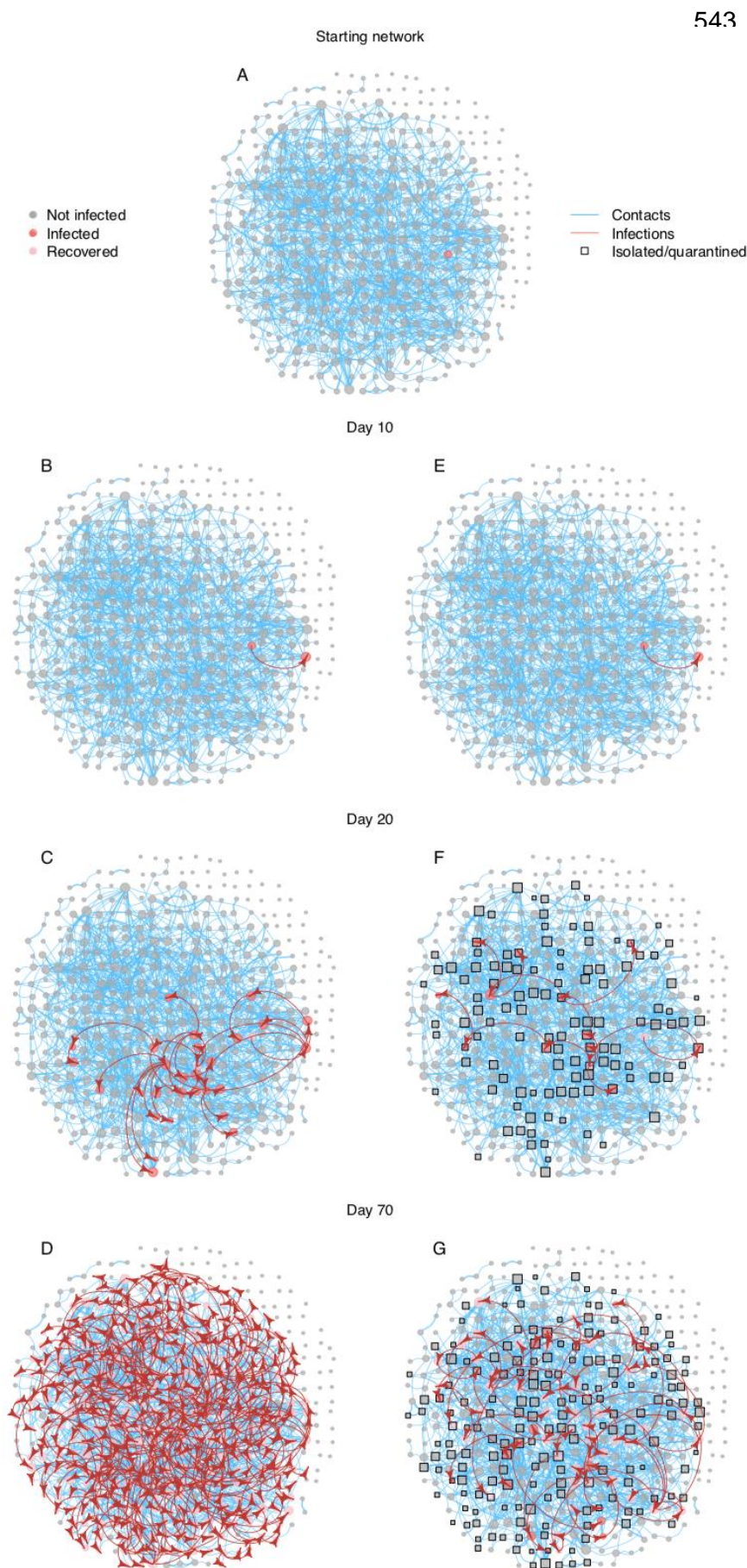
539 **Tables and Figures**

540

541 **Table 1** Policy summary

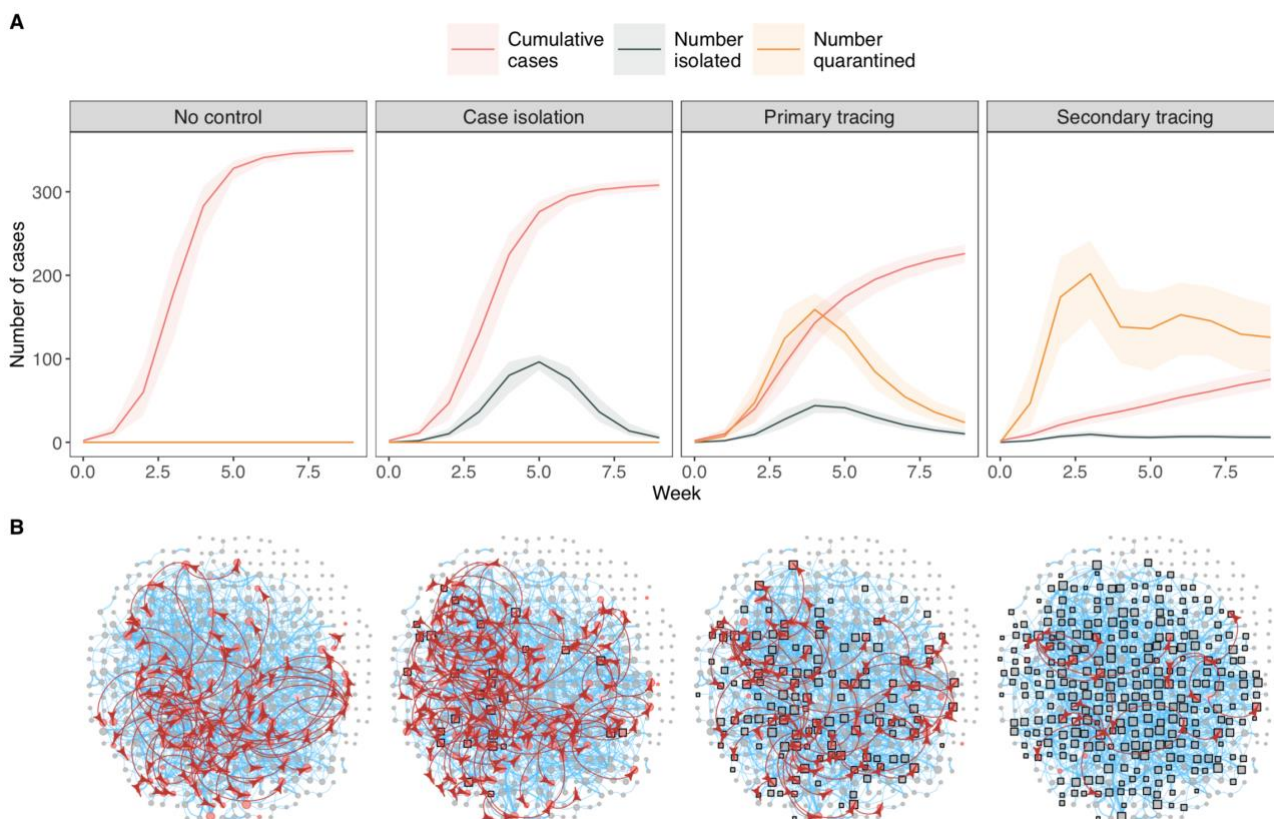
Background	Understanding how isolation, contact tracing and other non-pharmaceutical interventions can combine effectively and efficiently is crucial to COVID-19 control. Such interventions are likely to depend on contact patterns within a population. We developed an epidemic model that simulates COVID-19 outbreaks in a real-world network, and assess the impact of a range of testing, isolation, quarantine and contact tracing strategies for controlling these outbreaks.
Main findings and limitations	We found that isolation, contact tracing and quarantine reduced simulated outbreak size in our local-scale network. Tracing and quarantining contacts of contacts was more effective, but required large numbers of individuals are required to be quarantined. This strategy is therefore often similar to introducing a 'local lockdown'. Testing and releasing quarantined individuals reduced the numbers quarantined, but also the effectiveness of control measures. Combining physical distancing with contact tracing resulted in reduced outbreak size, with fewer individuals required to quarantine. A major limitation of this study is that it is based on pre-COVID-19 data from a sample of individuals from a single town; more data are therefore needed to fully understand potential outbreak dynamics in other settings.
Policy implications	Our findings suggest that effective contact tracing measures may require large numbers of people in a community to be quarantined, with individual-level tracing resulting in scenarios equivalent to broader localised lockdowns. Targeted tracing and quarantine strategies may therefore be more efficient when combined with other control measures such as physical distancing.

542



543

Figure 1 Illustration of the Haslemere network with epidemic simulation predictions. **A** The social network of 468 individuals (grey nodes) with 1257 social links (blue edges) weighted by 1616 daily contacts (edge thickness) and a single starting infector (red). Subsequent panels show progression of the COVID-19 epidemic under the no intervention (**B,C,D**) and the secondary contact tracing (**E,F,G**) scenarios. Red arrows show an infection route, and squares show isolated/quarantined individuals.

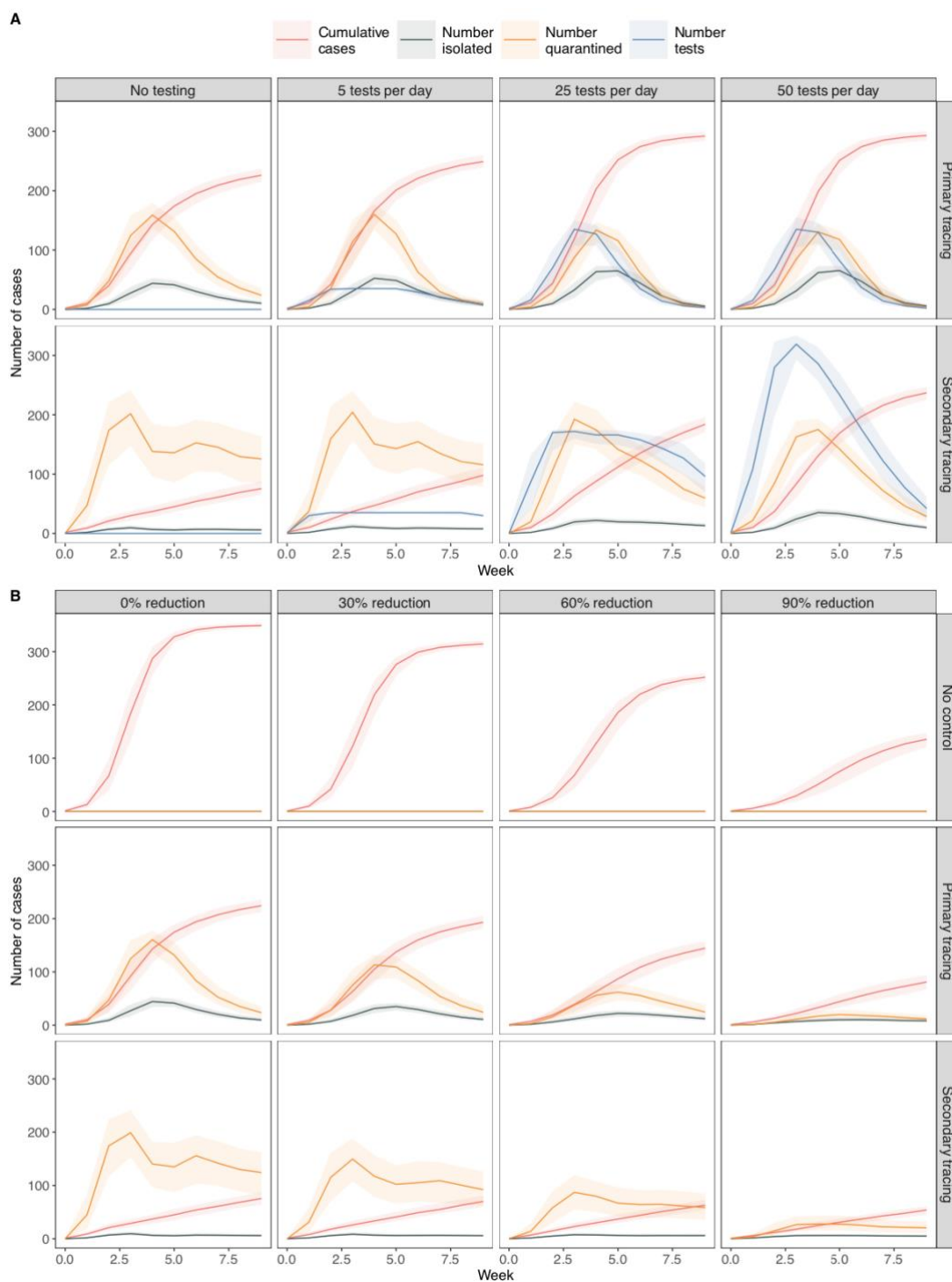


566

567 **Figure 2** Epidemic model predictions of outbreak size and number of people isolated/quarantined
568 under different non-pharmaceutical intervention scenarios in the Haslemere network. **A** cumulative
569 number of cases, number of people isolated, and number of people quarantined at a given point in
570 time under each scenario. Lines and shaded areas represent median and interquartile range from
571 1000 simulations. **B** Example networks from a single simulation of each scenario at day 20 of the
572 outbreak. See figure 1 for network details.

573

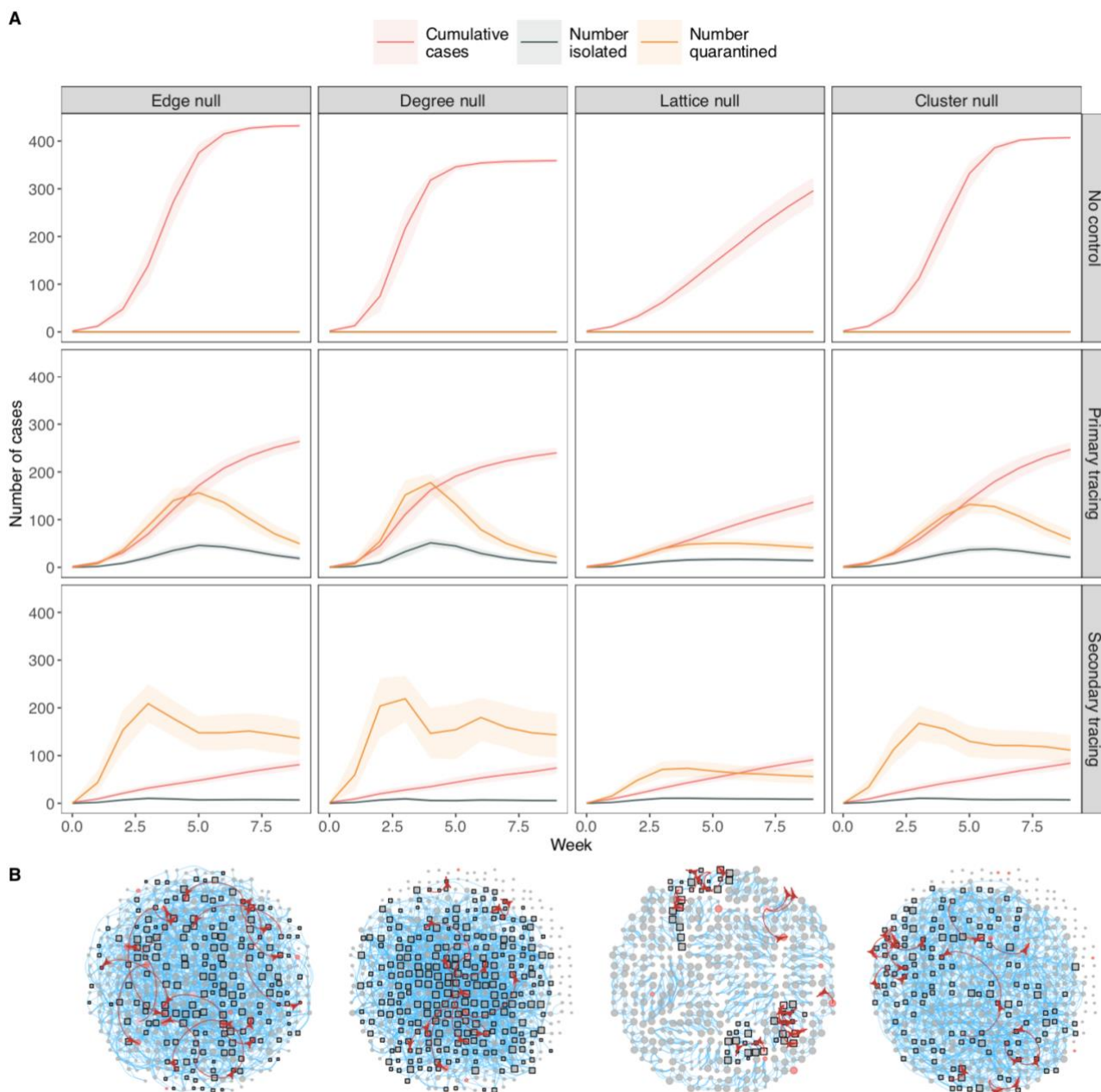
574



575

576 **Figure 3 A** Epidemic model simulations of outbreak size and number of people isolated and
 577 quarantined under **A** different levels of testing and **B** physical distancing in the Haslemere network.

578 In **A**, Tests are plotted per week rather than per day for visualisation purposes. In **B** The
 579 percentage reduction refers to the number of 'weak links' removed from the networks (see
 580 methods). Lines and shaded areas represent median and interquartile range from 1000
 581 simulations.



582

583 **Figure 4 A** Epidemic model simulations of outbreak size and number of people isolated and
 584 quarantined under different null-network permutations based on the Haslemere network (see
 585 methods for details). Lines and shaded areas represent median and interquartile range from 1000
 586 simulations. **B** Example networks showing an infection simulation (with secondary contact tracing,
 587 after 20 days) on each null network. See Figure 1 for network details.